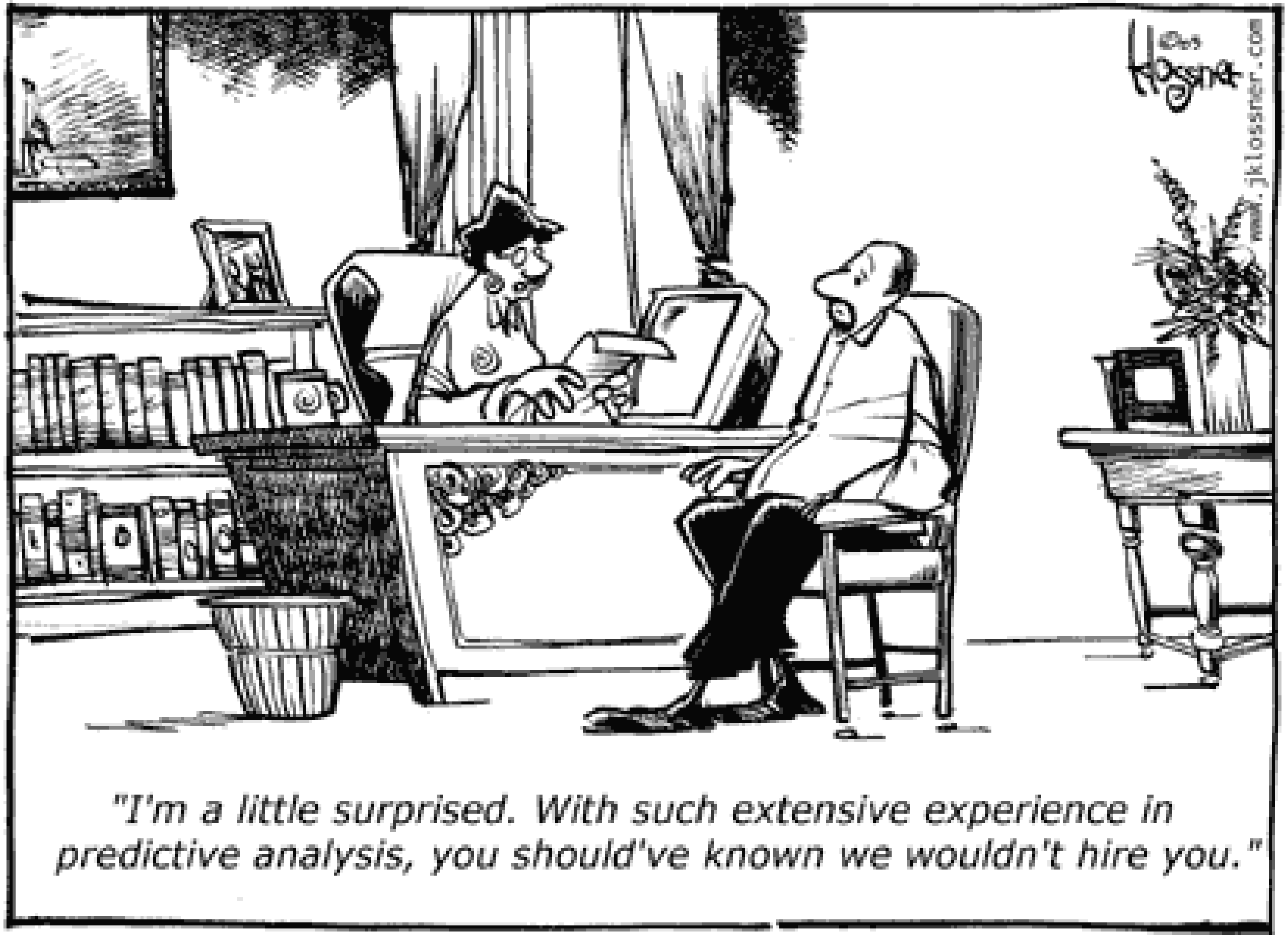
Making Predictions at the Edge Of Chaos

Ashok N. Srivastava, Ph.D.
Principal Investigator, IVHM Project
Group Leader, Intelligent Data Understanding Group

Santanu Das, Ph.D.
Arizona State University
NASA Ames Research Center

June 2007

Some Predictions are Difficult



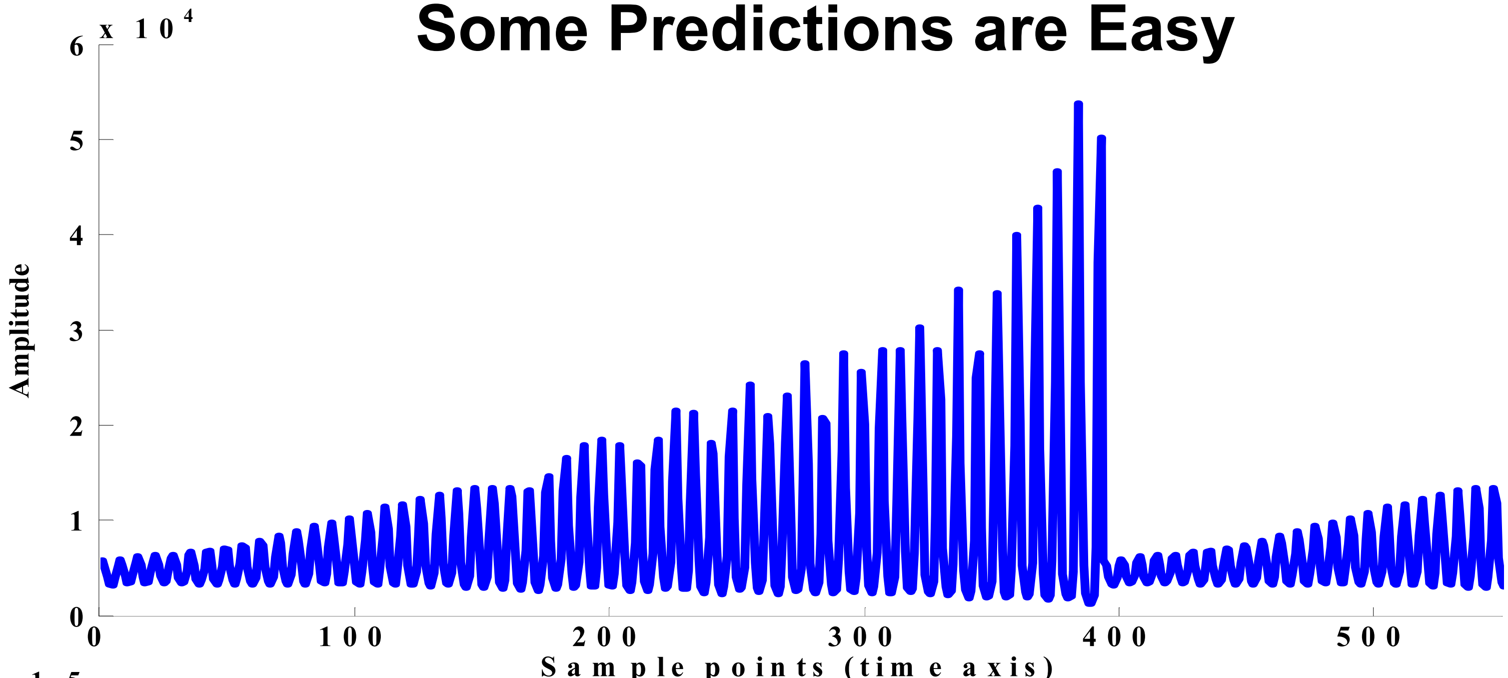
"I'm a little surprised. With such extensive experience in predictive analysis, you should've known we wouldn't hire you."

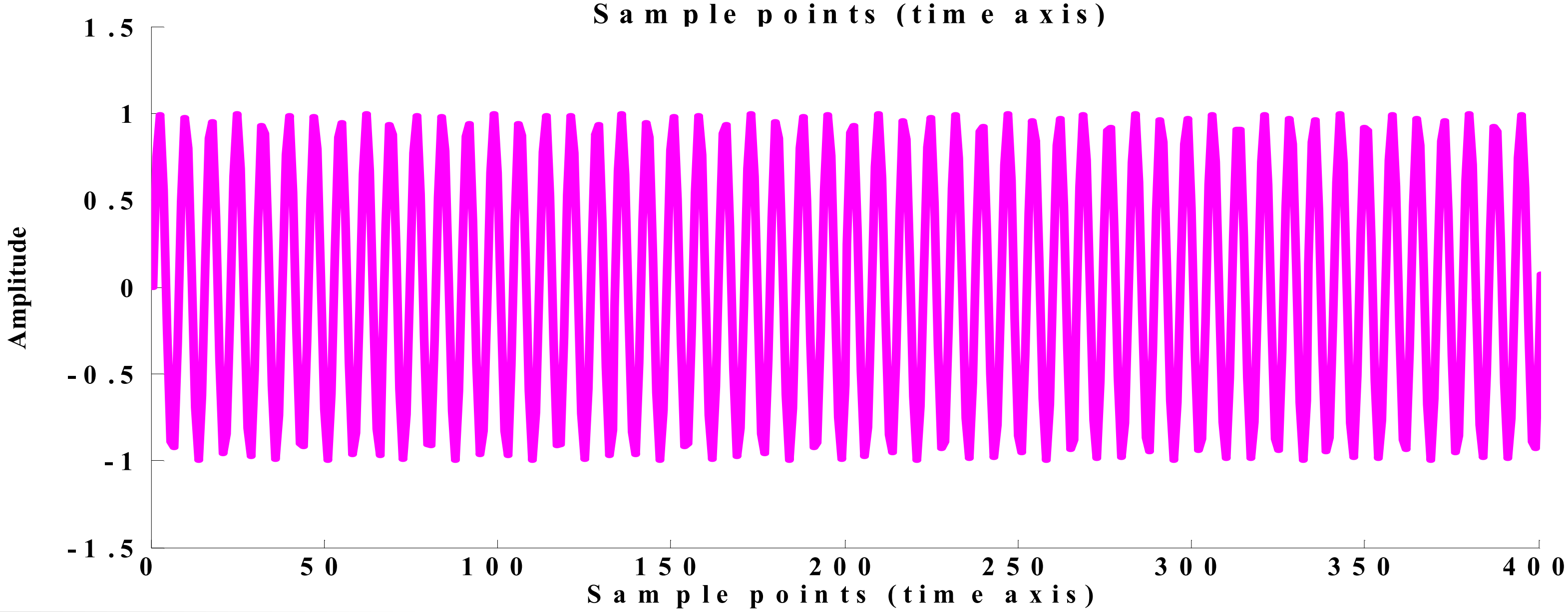
One of Leibniz's Views on Prediction

"If someone could have a <u>sufficient insight</u> into the inner parts of things, and in addition had remembrance and intelligence enough to consider all the circumstances and to take them into account, he would be a prophet and would see the future in the present as in a mirror."

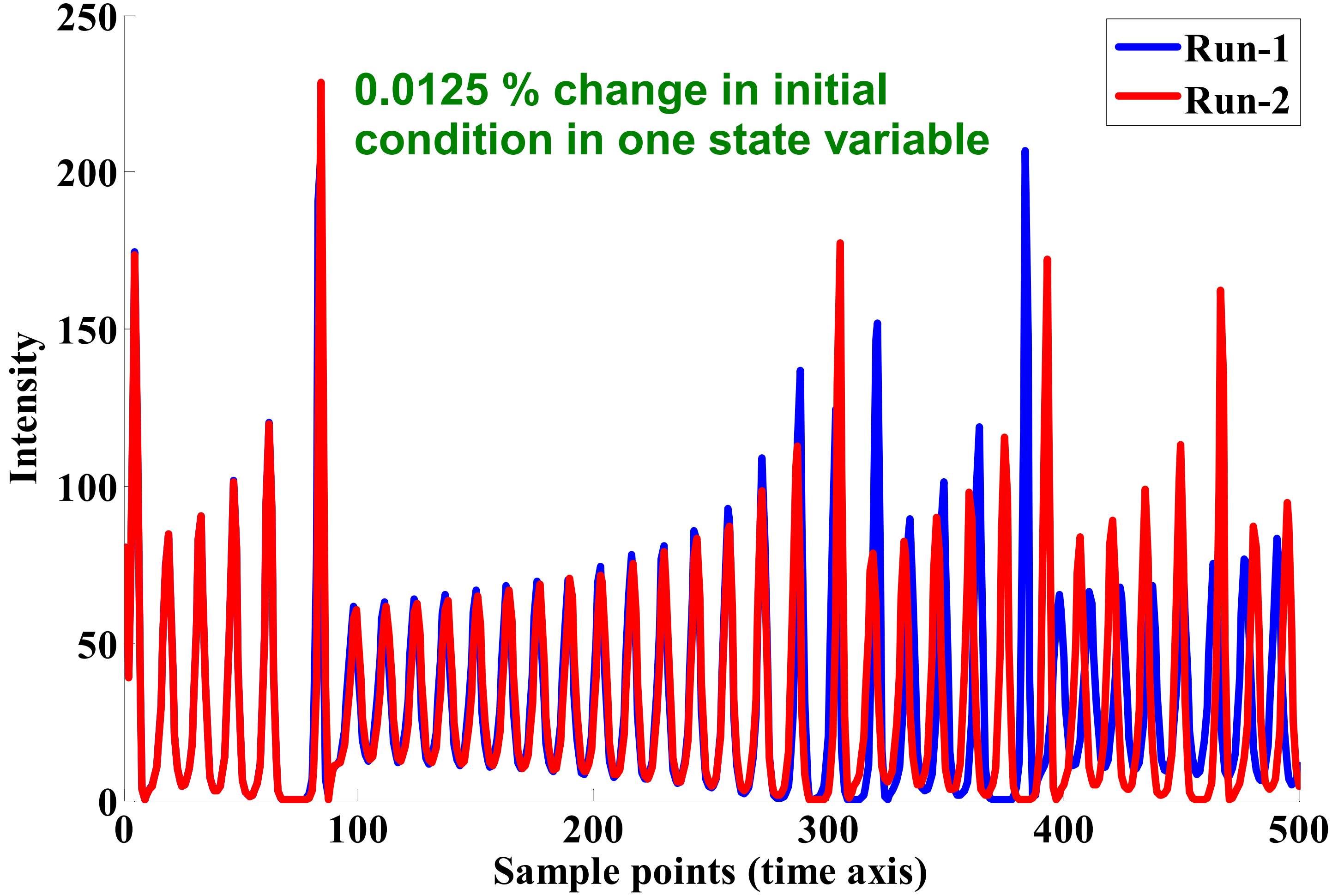
From ChaosBook.org and Wikipedia

Some Predictions are Easy





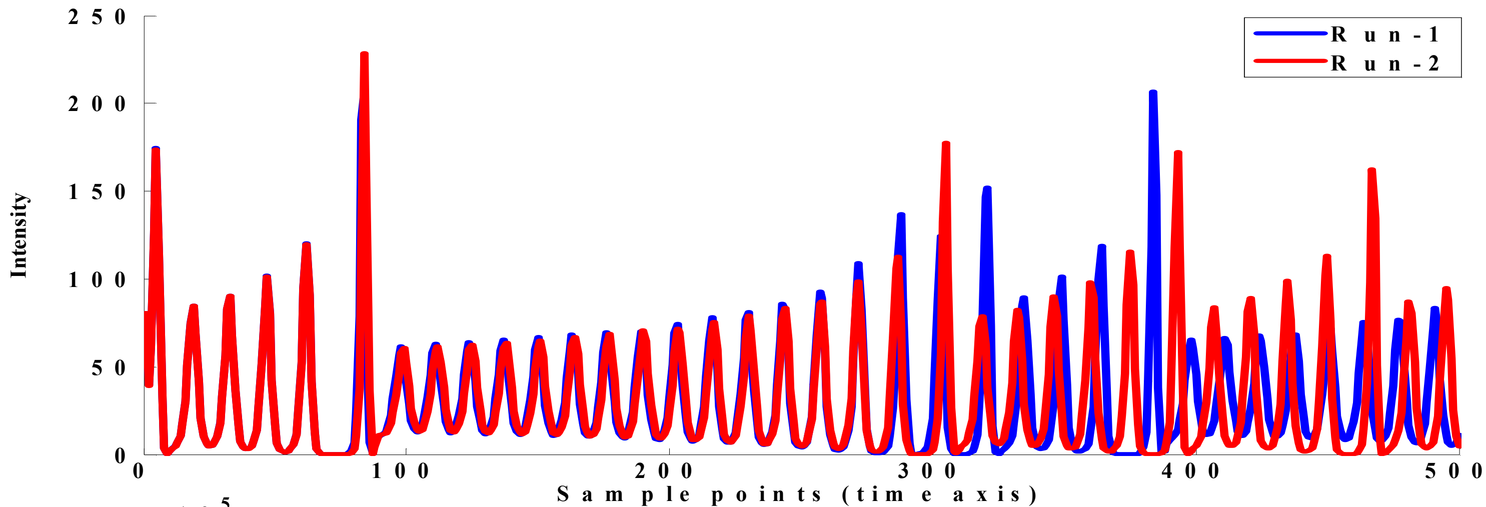
Lyapunov Exponents and the Limits on Predictability

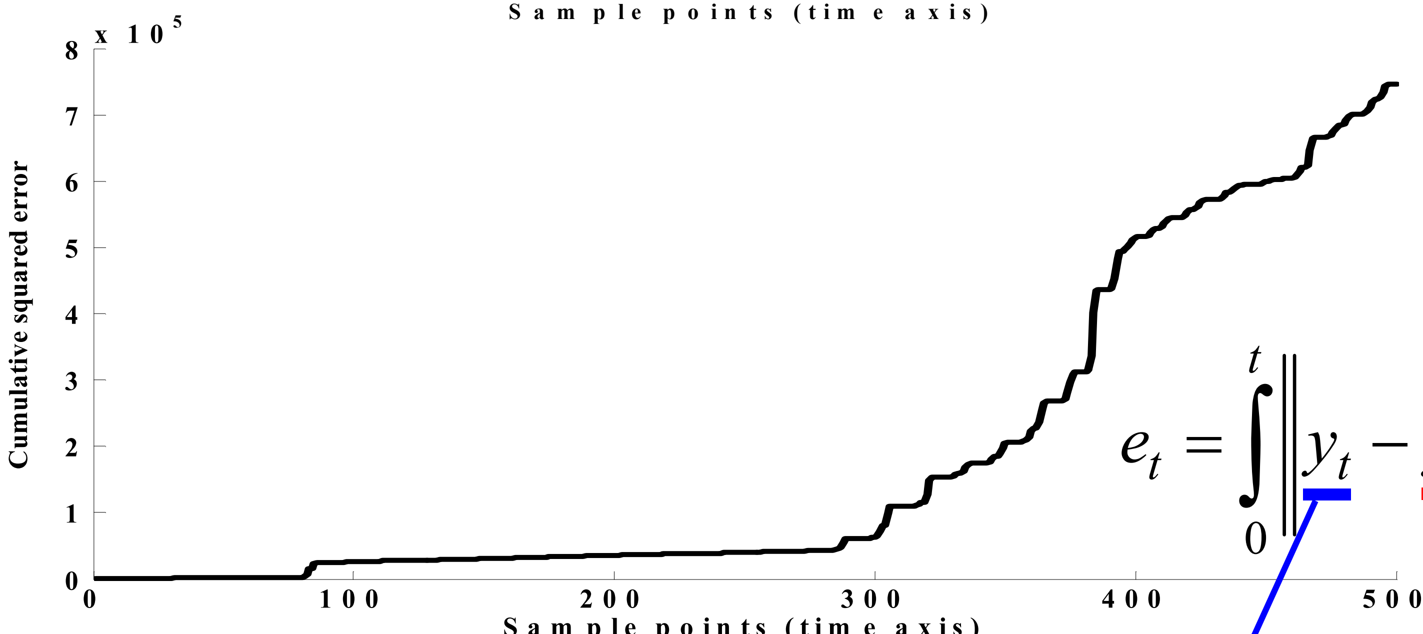


$$\delta Z(t) \approx e^{\lambda t} \left| \delta Z_0 \right|$$

$$\lambda = \lim_{t \to \infty} \frac{1}{t} \ln \frac{\left| \delta Z(t) \right|}{\left| \delta Z_0 \right|}$$

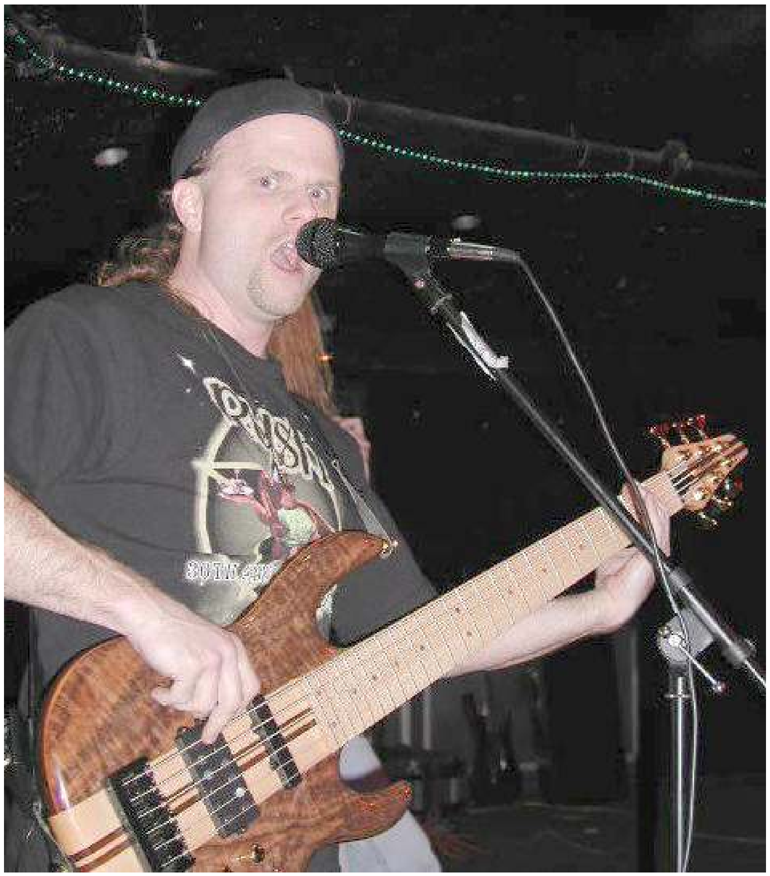
The Edge of Chaos





$$e_t = \int_0^t \left\| \underbrace{y_t}_{\text{Run-1}} - \underbrace{\hat{y}_t}_{\text{Run-2}} \right\|^2 dt$$

Edge Of Chaos



Extreme Shred Metal

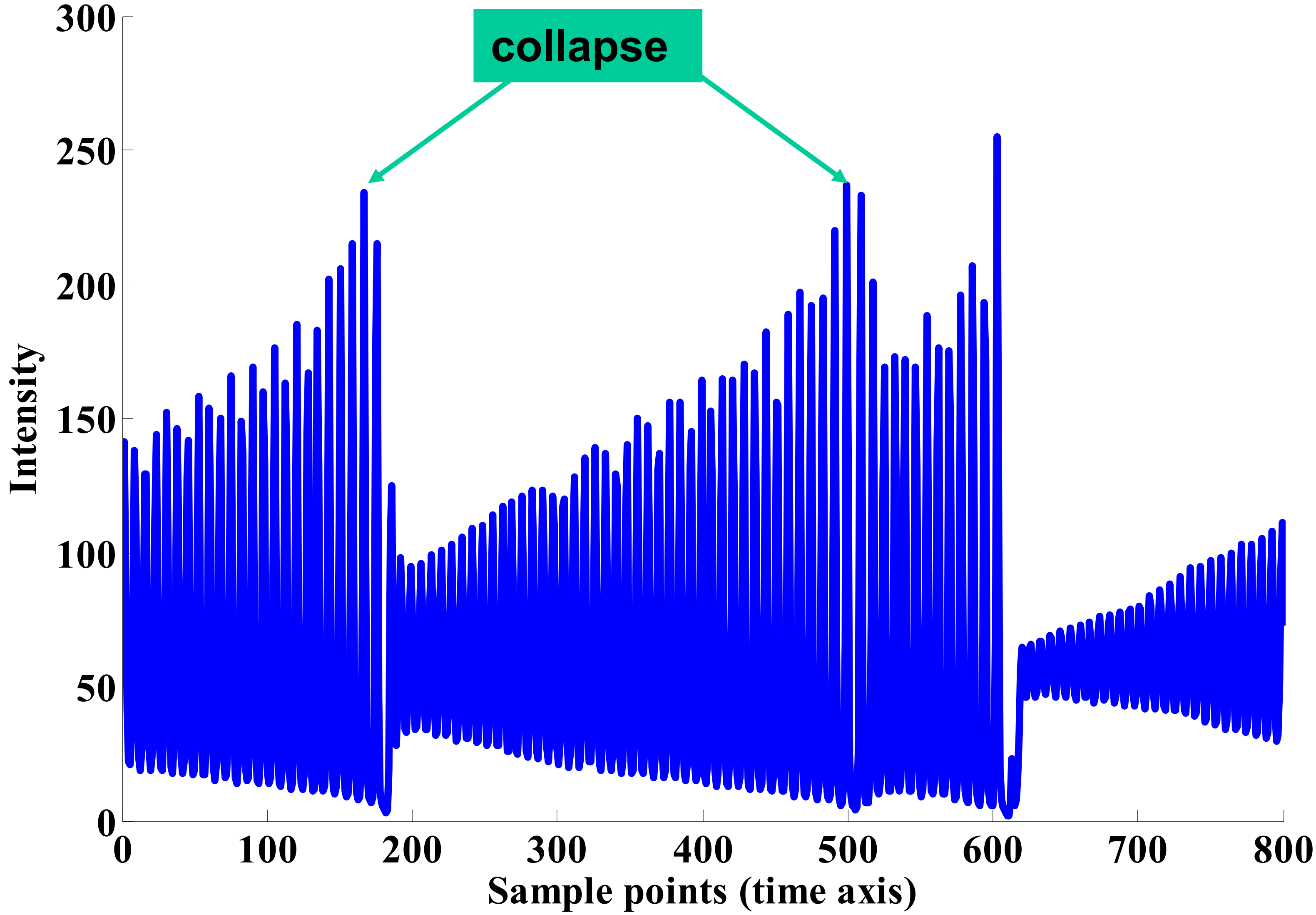
www.edgeofchaos.us

Applications to Laser Systems

- Develop a set of algorithms for prognostics using data from a well-studied ammonium laser system that has chaotic behavior.
- Predict the future dynamics of this system
- Generate a signal that represents the confidence in the prediction.

The Need for Prognostics

Observed Laser Intensity



NH_3 Laser Model

One can approximate the dynamical behavior of the laser using ideas from nonlinear dynamical systems.

Lorenz-Hankel Model

$$\frac{dx}{dt} = \sigma(y - x)$$

$$\frac{dy}{dt} = rx - y - xz$$

$$\frac{dz}{dt} = xy - bz$$

Nonlinear terms

r, σ, b ***Control Parameters***

$\boxed{r = 15}$

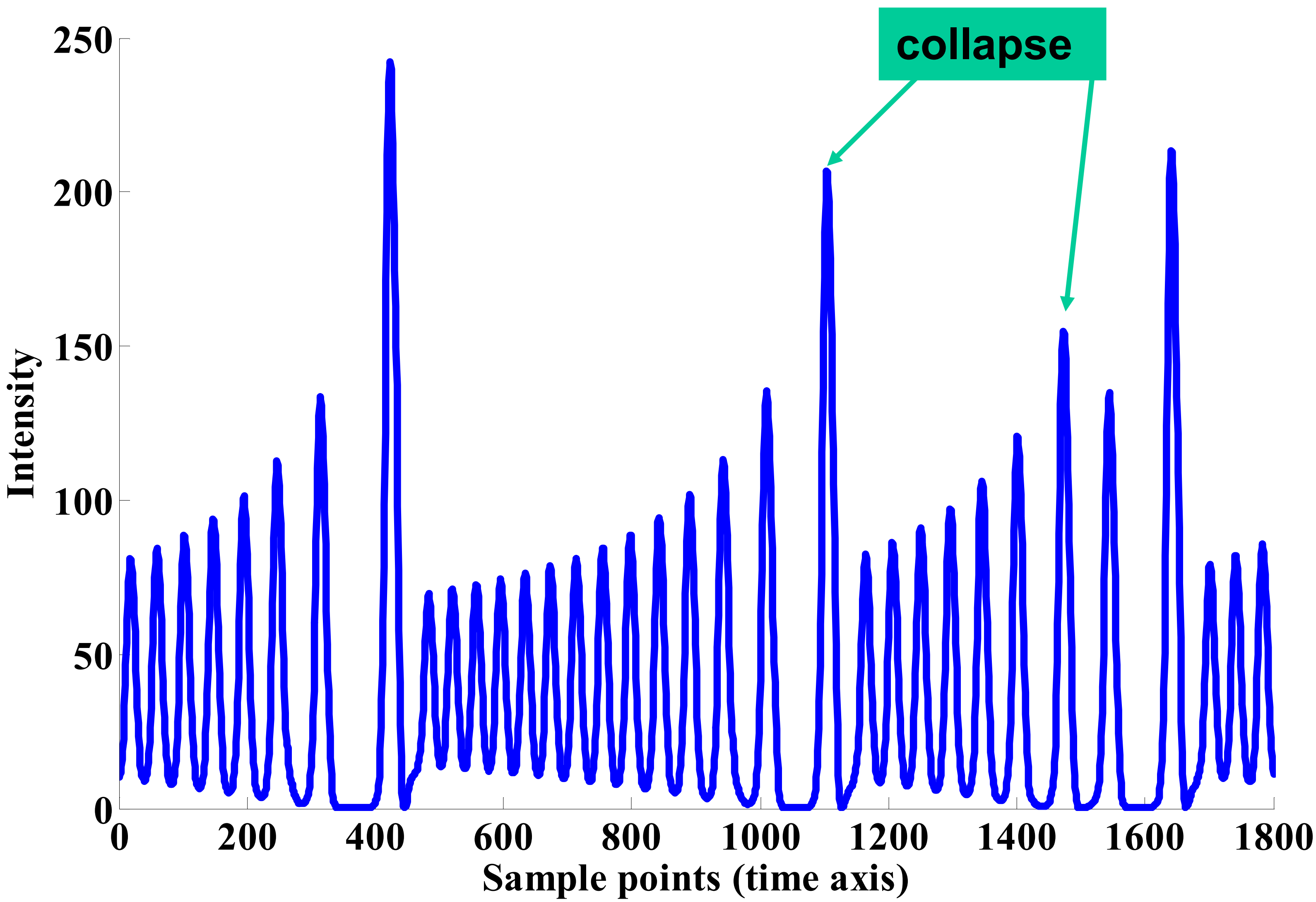
$\boxed{\sigma = 0.2}$

$\boxed{b = 0.25}$

$\boxed{q < 0.05}$

The values of sigma, r, b and q determine the nature of the chaotic attractor.

Lorenz-Hankel Model



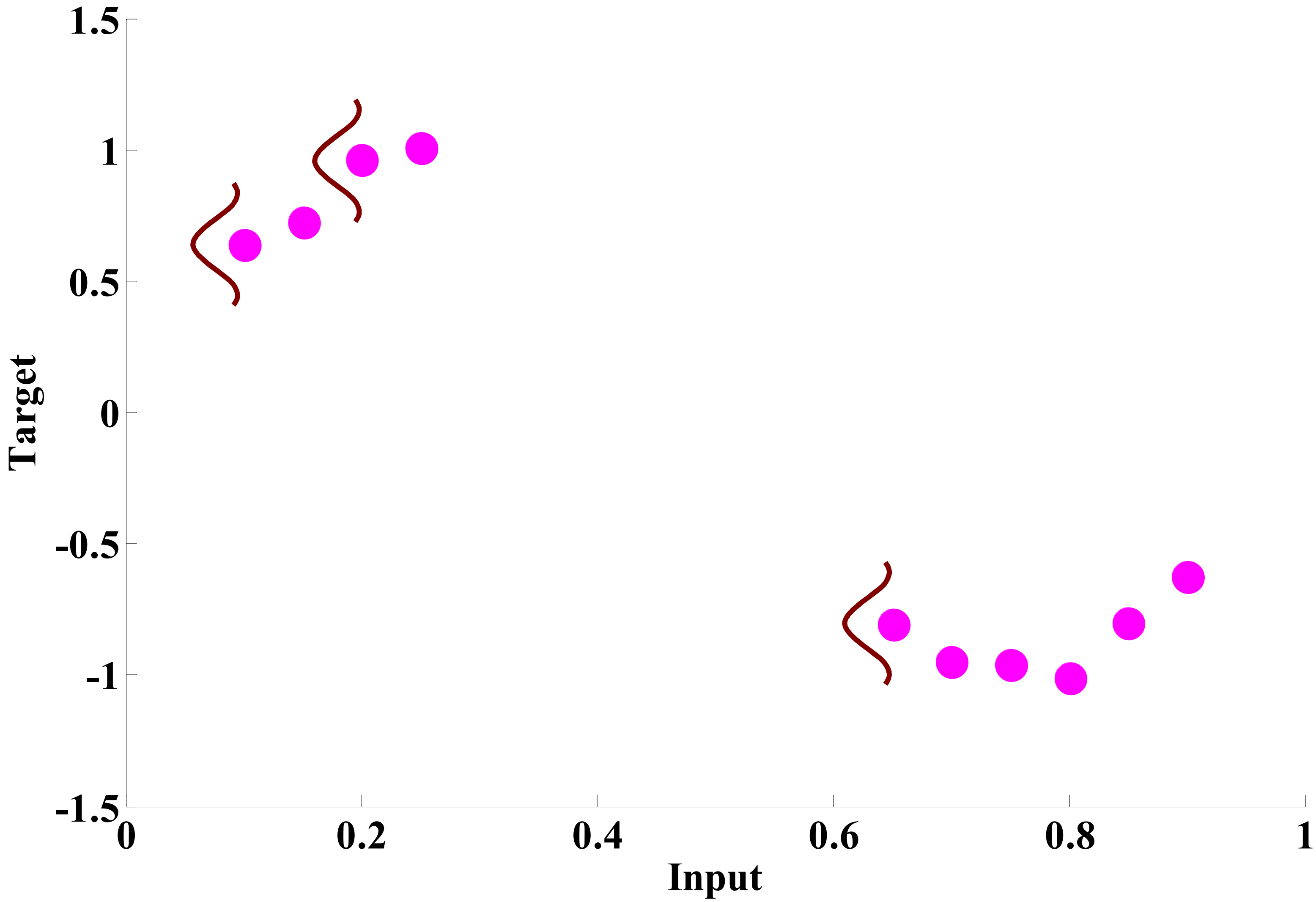
Gaussian Process (GP)

Any random function is a GP, if $\{f(x_1), f(x_2), \ldots, f(x_n)\}$ is a random vector which is normally distributed for all $x_1, x_2, \ldots, x_n$.

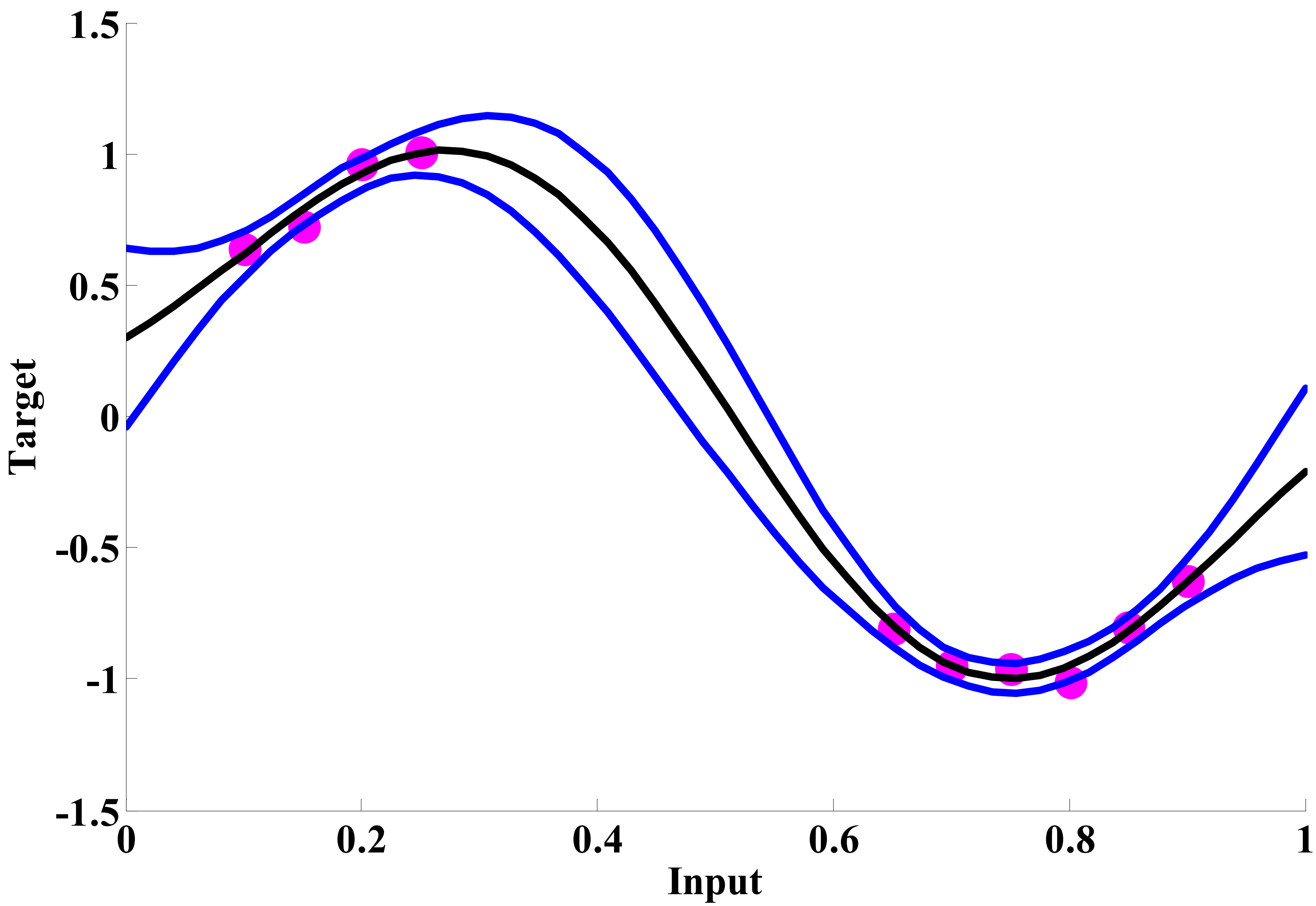
$$p\left(f(x_1), f(x_2), \ldots, f(x_n) \middle| x_1, x_2, \ldots, x_n\right) = N\left(m(x), C(x_i, x_j)\right)$$

Each function is characterized by its mean $\underset{n \times 1}{m(x)}$ and variance $\underset{n \times n}{C(x_i, x_j)}$

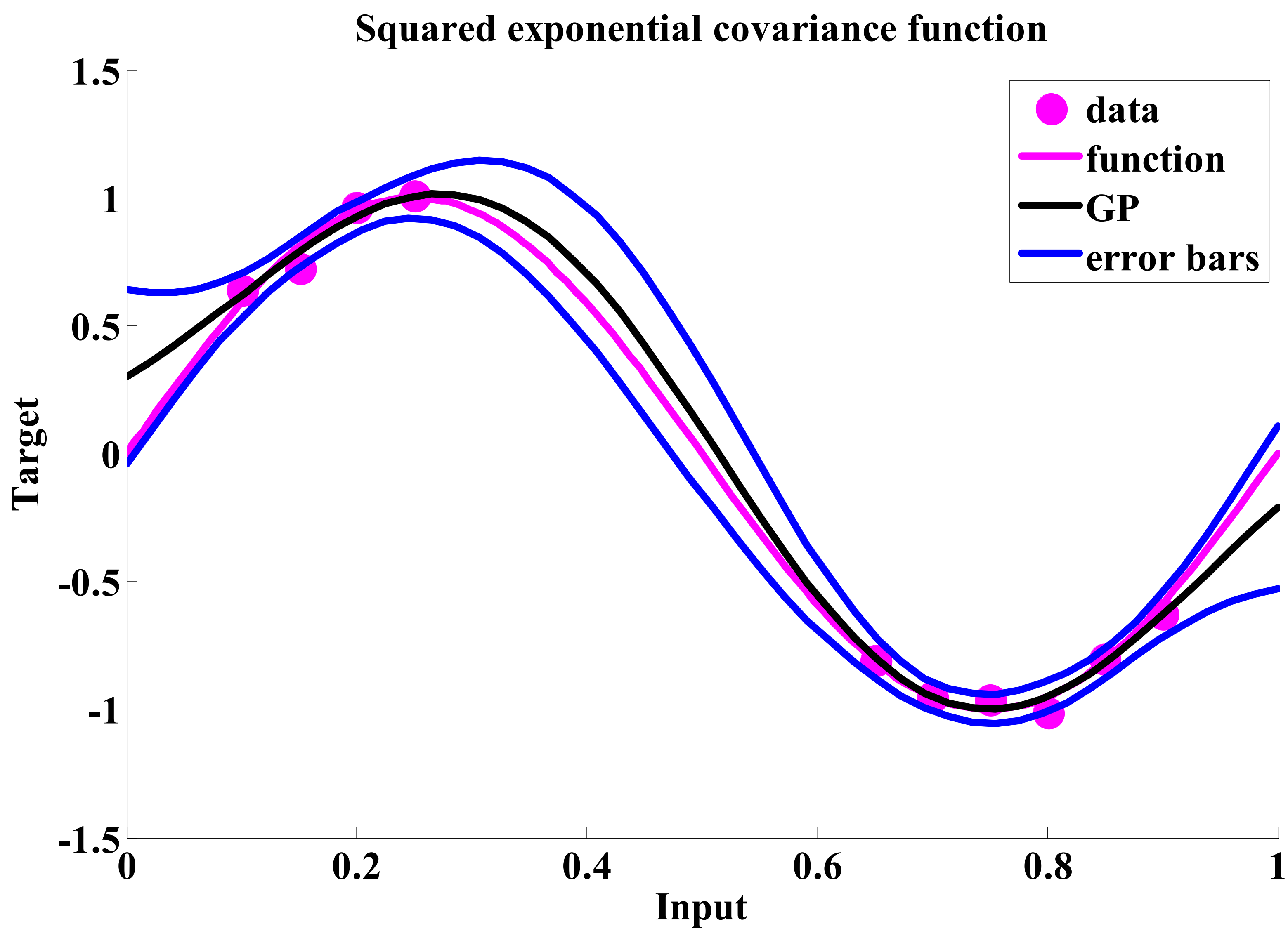
Gaussian Process Regression Chooses the Best Function to Explain a Data Set



The Covariance Function Determines the Fit



Gaussian Process Regression Example



Example Covariance Functions

Example: When the process is stationary

Assumption: function *smooth* & *continuous*

Mean $m = const.$ (here zero)

Input dimension

$$C(x_i, x_j) = \theta \exp\left(-\frac{1}{2} \sum_{k=1}^{D} \frac{(x_i^k - x_j^k)^2}{\sigma_k^2} \right)$$

Gaussian

Hyperparameters

$$C(x_i, x_j) = < x_i, x_j >^2$$

Quadratic

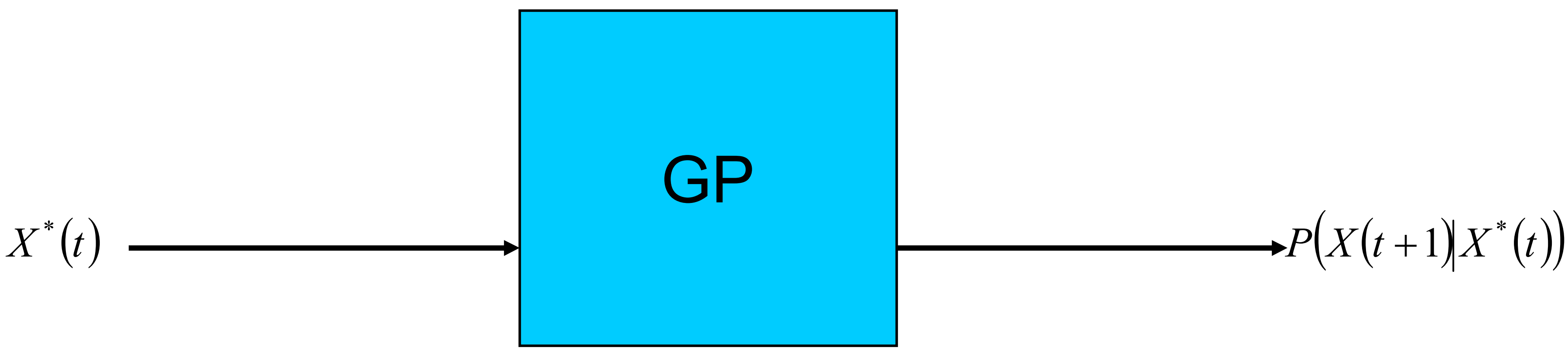
Approach

- Using delay coordinate embedding (and thus Takens' Theorem) we build a Gaussian Process Regression (GPR) to predict:

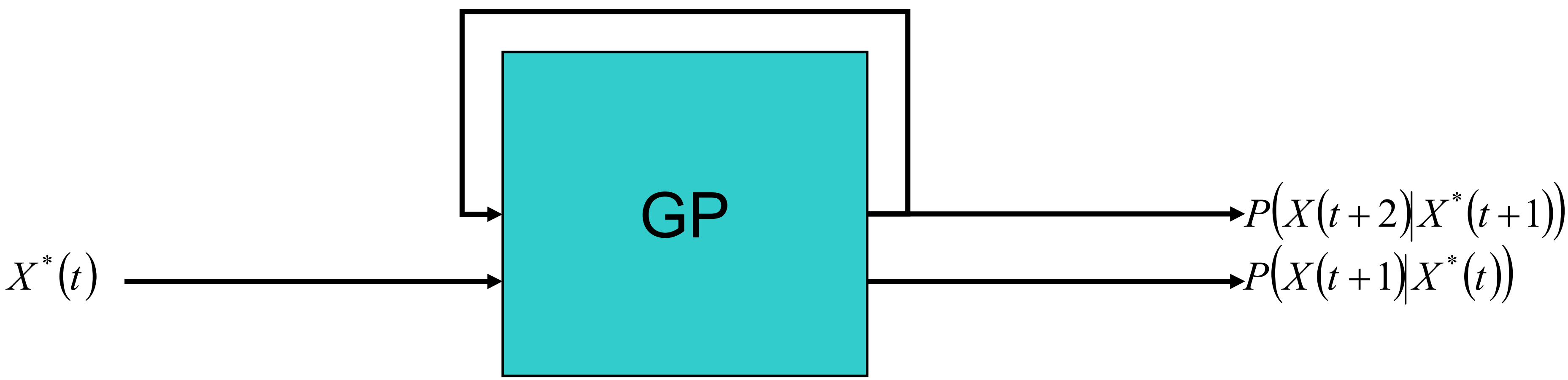
$$P\left(X(t+1) \middle| X(t), X(t-1), \ldots., X(t-\underbrace{d}_{\textit{Embedding dimension}})\right) = P\left(X(t+1) \middle| X^*(t)\right)$$

- Once this distribution is known, we can make predictions through iterating the distribution.

One Step Ahead Predictions



Iterated Predictions

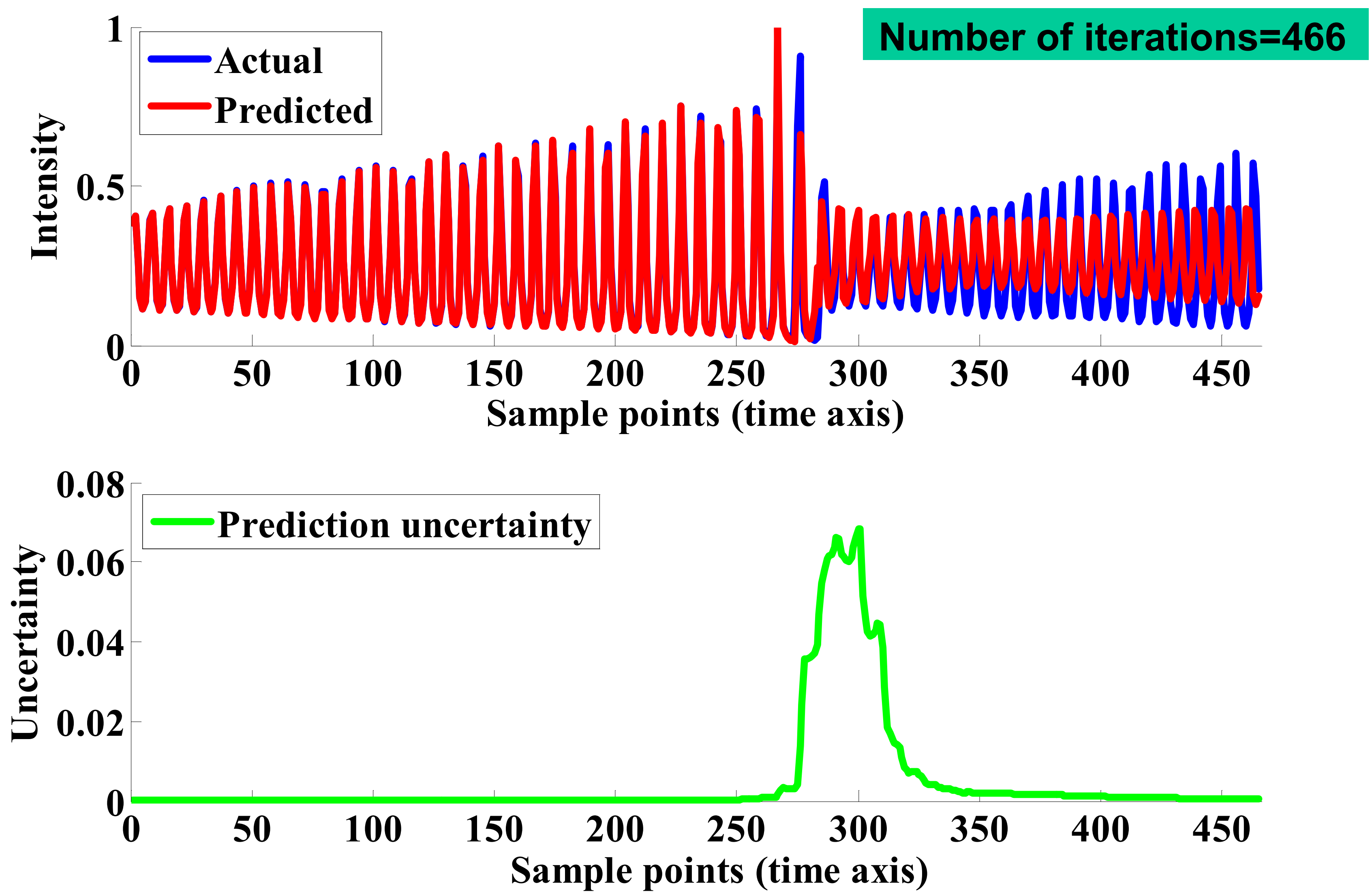


i.e., we feed the output of the model into its input to make a prediction of

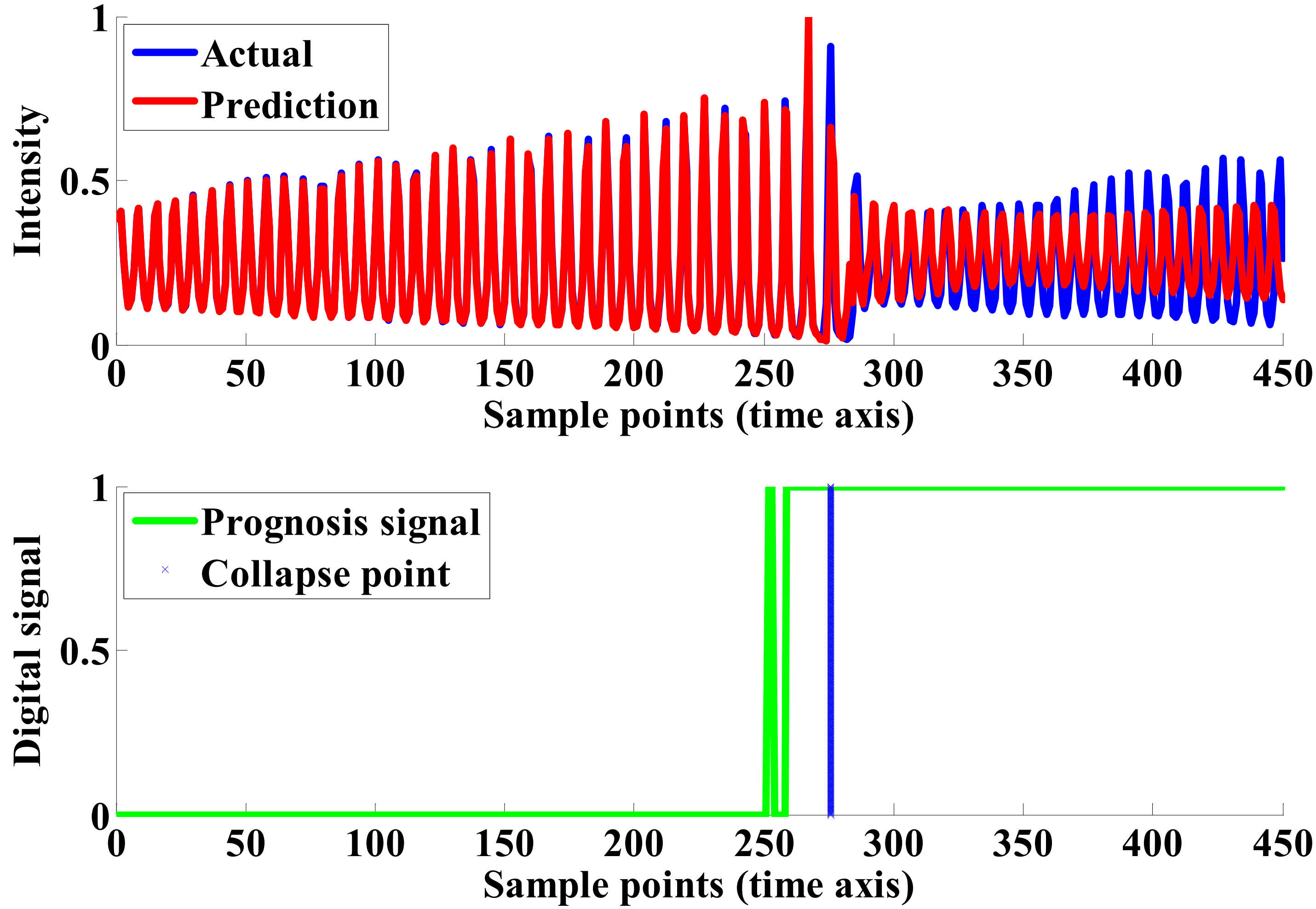
$$P\big(X(t+2)\big|\big[P\big(\underline{X(t+1)}, X(t), X(t-1), \ldots., X(t-d+1)\big)\big]\big) = P\big(X(t+2)\big|X^*(t+1)\big)$$

From past prediction iteration

Iterated Gaussian Process Predictions

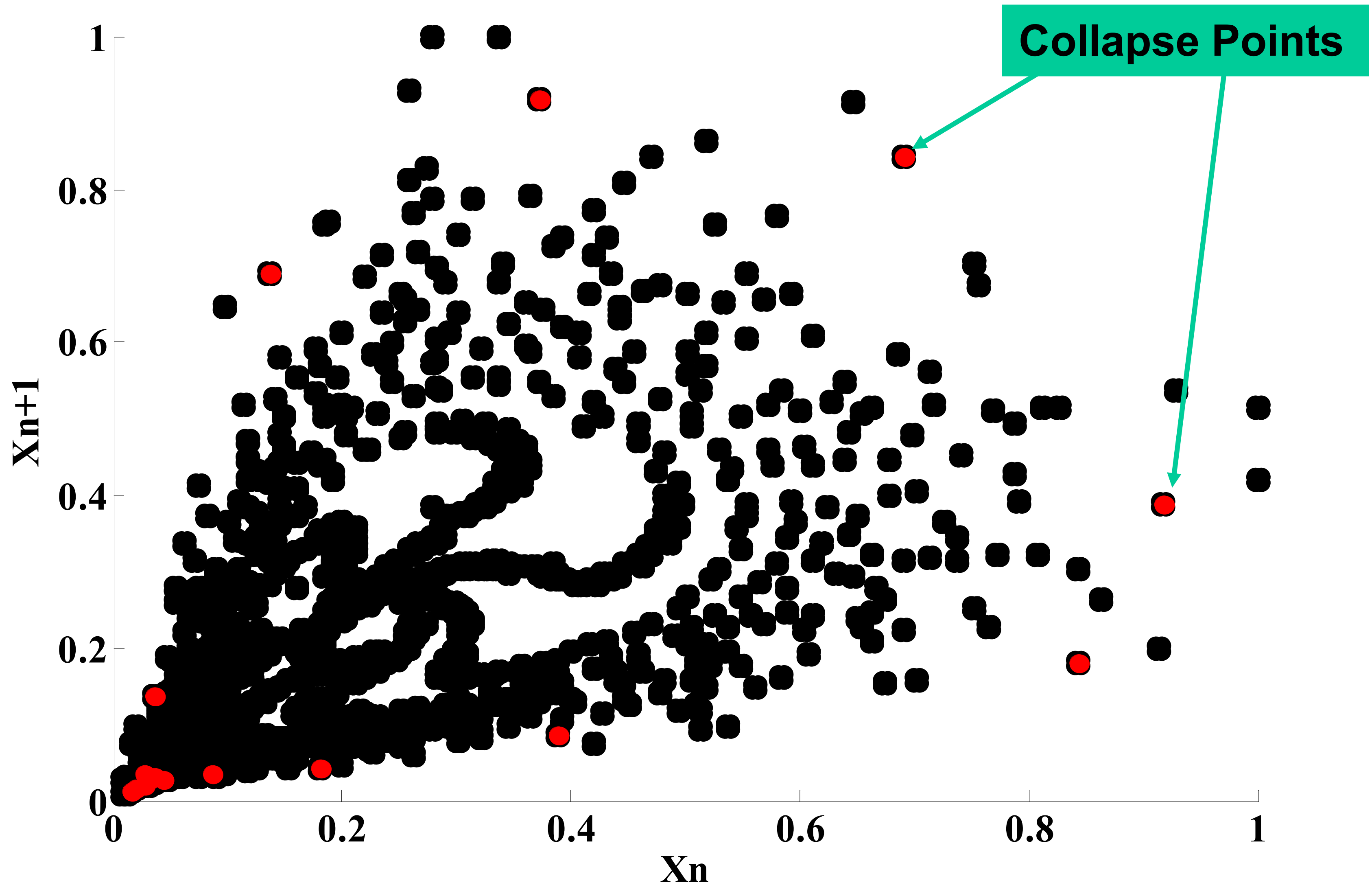


Prognostic Signal

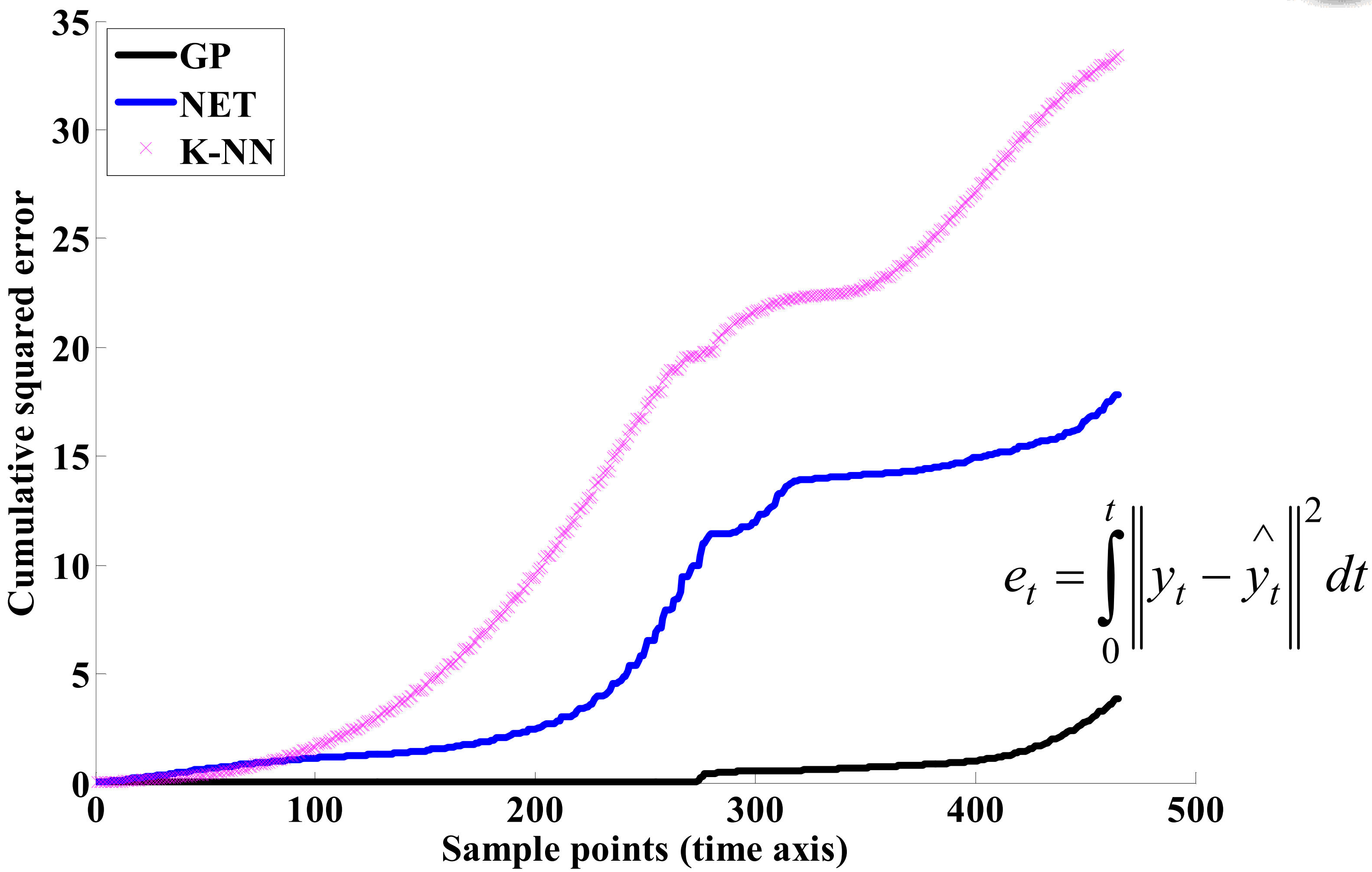


Prediction signal leads the actual collapse point by 24 sample points

Phase Portrait



Cumulative Error Curves



Threshold

Cumulative squared error <= 1

GP	Bagged NN	K-NN
397	84	79

Results

- We have shown that we can make iterated forecasts and detect a precursor to the sudden drop in intensity using kernel methods.
- We can generate a meaningful measure of prediction certainty.
- This quantity seems to indicate substantial increases in uncertainty near the collapse.

Structural Application

- Presence of partially closed cracks in objects can be identified using an ultrasonic technique. (Ref: K Yamanaka)
- Interaction of high amplitude ultrasonic waves with closed cracks generate subharmonic components.

Characteristic length of crack plane

$$m\ddot{x} + \gamma\dot{x} + k\left(x - \underbrace{x_s}_{\text{Equilibrium Position}}\right) = f_0\left[\kappa\left(\frac{\sigma}{x - a\sin\omega t}\right)^{M} - \left(\frac{\sigma}{\underbrace{x - a\sin\omega t}_{\text{COD}}}\right)^{N}\right]$$

Equilibrium Position

Repulsive force parameters

Attractive force parameters

- The vibration of the Crack Opening Displacement (COD) exhibits chaotic behavior if:

$$x_s = 10\sigma, f_0 = 15, m = 1, \gamma = 0.5, k = 0.2, \omega = 1, x(t_0) = 1.8\sigma \qquad \boxed{a = 8\sigma}$$

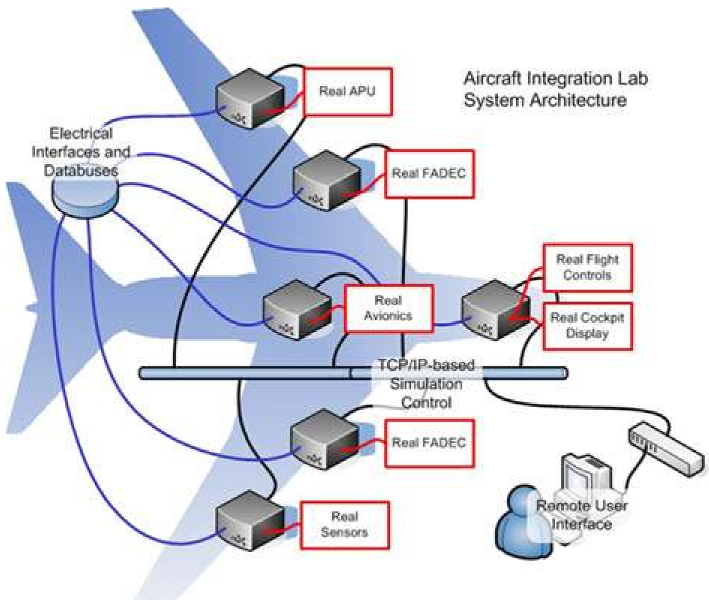
Further Work

- Understanding the limits of predictability for these systems and for other non-time based predictive models
- Significant testing with respect to forecast variability and quality of precursor detection.
- Analysis of forecast horizon and comparison with Lyapunov Time.
- Test methods on data from aircraft systems.

IVHM Data Mining Lab

Mission of the IVHM Data Mining Lab

The lab enables the dissemination of Integrated Vehicle Health Management data, algorithms, and results to the public. It will serve as a national asset for research and development of discovery algorithms for detection, diagnosis, prognosis, and prediction for NASA missions.



Features of the IVHM Data Mining Lab

Datasets

- Propulsion, structures, simulation and modeling
- ADAPT Lab
- Icing
- Electrical Power Systems
- Systems Analysis
- Flight and subscale systems
- Fleet-wide data
- Multi-carrier data

Open Source

- Code
- Papers
- Generation of an IVHM community

Selected Discovery Tools

- Inductive Monitoring System (IMS) – cluster-based anomaly detection
- Mariana – Text classification algorithm
- Orca – Distance-based outlier detection
- ReADS – Recurring anomaly detection system for text
- sequenceMiner – anomaly detection for discrete state and mode changes in massive data sets.

Key Research Issues Addressed in the IVHM Data Mining Lab

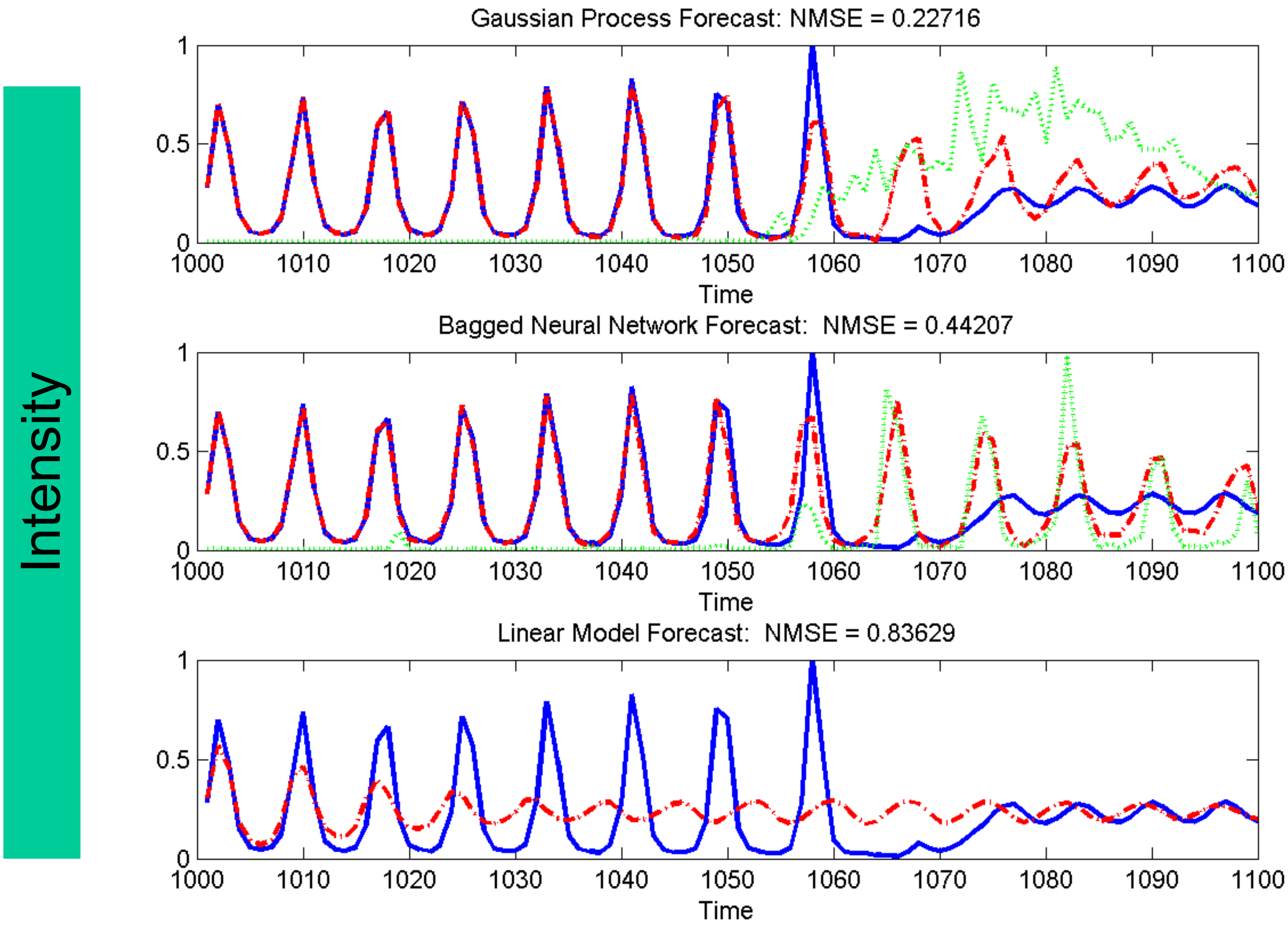
- Real-time anomaly detection
- Model-free prediction methods
- Hybrid methods that combine discrete and continuous data
- Distributed and privacy-preserving data mining
- Analysis of integrated systems

Appendix

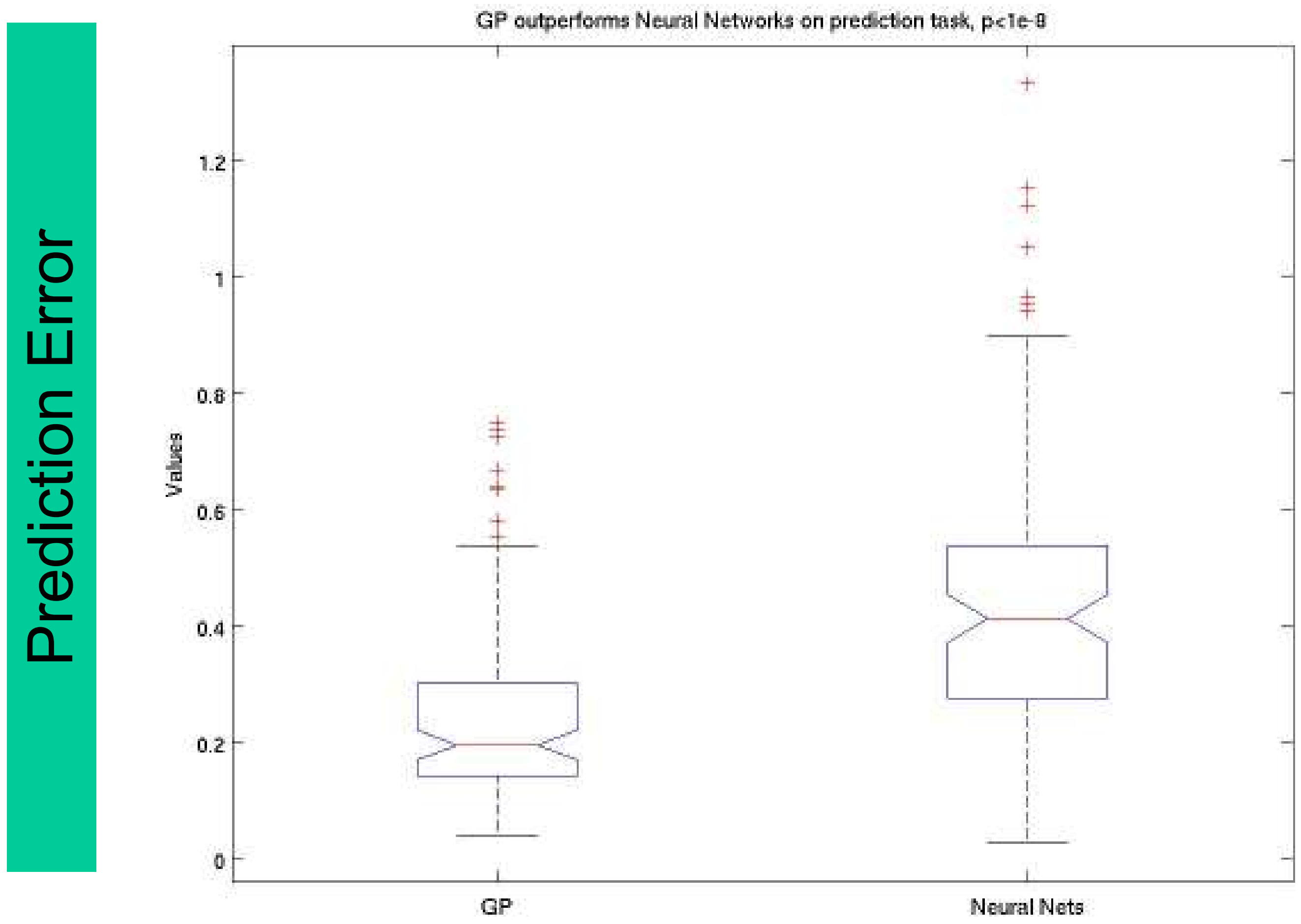
NH_3 Laser Phenomena

- The laser undergoes periods of buildup of intensity followed by a sudden collapse in intensity.
- Sometimes the collapse is significant, and other times it is relatively small.
- It is hard to predict what type of collapse will occur (i.e., it is a chaotic process).

Comparison


Intensity
Gaussian Process Forecast: NMSE = 0.22716
1
0.5
0
1000
1010
1020
1030
1040
1050
1060
1070
1080
1090
1100
Time
Bagged Neural Network Forecast: NMSE = 0.44207
1
0.5
0
1000
1010
1020
1030
1040
1050
1060
1070
1080
1090
1100
Time
Linear Model Forecast: NMSE = 0.83629
1
0.5
0
1000
1010
1020
1030
1040
1050
1060
1070
1080
1090
1100
Time

Statistical Comparison of GP's and Neural Networks



K-step ahead forecasts

- We iterate the Gaussian Process K times to generate this time series.

- Performance comparison
  - » Bagged Neural Networks
  - » Linear Model

- Forecasting metric:
  normalized mean squared error

Method

- We address this problem using the theory of Gaussian Processes which assumes that any subset of data for a vector X is Gaussian distributed (from wikipedia).

$$\vec{\mathbf{X}}_{t_1,\ldots,t_k} = (\mathbf{X}_{t_1},\ldots,\mathbf{X}_{t_k})$$

Using characteristic functions of random variables, we can formulate the Gaussian property as follows: $\{X_t\}_{t \in T}$ is Gaussian if and only if for every finite set of indices $t_1, \ldots, t_k$ there are positive reals $\sigma_{l j}$ and reals μ_j such that

The numbers $\sigma_{l j}$ and μ_j can be shown to be the covariances and means of the variables in the process.

References

- A. S. Weigend and N. Gershenfeld, “Time Series Prediction: Forecasting the Future and Understanding the Past”, 1994
- Gaussian Process Regression, J.S. Taylor, 2002